

SEQUENCE LISTING

<110> SYRRX, INC.

<120> CRYSTALLIZATION OF CATHEPSIN S

<130> SYR-CATS-5002-C1

<140> Not Yet Assigned

<141> 2003-08-22

<150> US 60/405,423

<151> 2002-08-23

<160> 4

<170> PatentIn version 3.1

<210> 1

<211> 331

<212> PRT

<213> Homo sapiens

<220>

<221> Amino acid sequence for full-length human wild type Cathepsin S

<222> (1)..(331)

<223>

<220>

<221> Amino acid sequence for full-length human wild type Cathepsin S

<222> (1)..(331)

<223> Residues 115-331 comprise the catalytic domain

<300>

<308> AF230097

<309> 2002-04-08

<313> (1)..(331)

<400> 1

Met Lys Arg Leu Val Cys Val Leu Leu Val Cys Ser Ser Ala Val Ala
1 5 10 15

Gln Leu His Lys Asp Pro Thr Leu Asp His His Trp His Leu Trp Lys
20 25 30

Lys Thr Tyr Gly Lys Gln Tyr Lys Glu Lys Asn Glu Glu Ala Val Arg
35 40 45

Arg Leu Ile Trp Glu Lys Asn Leu Lys Phe Val Met Leu His Asn Leu
50 55 60

Glu His Ser Met Gly Met His Ser Tyr Asp Leu Gly Met Asn His Leu
65 70 75 80

Gly Asp Met Thr Ser Glu Glu Val Met Ser Leu Met Ser Ser Leu Arg
85 90 95

Val Pro Ser Gln Trp Gln Arg Asn Ile Thr Tyr Lys Ser Asn Pro Asn
100 105 110

Arg Ile Leu Pro Asp Ser Val Asp Trp Arg Glu Lys Gly Cys Val Thr
115 120 125

Glu Val Lys Tyr Gln Gly Ser Cys Gly Ala Cys Trp Ala Phe Ser Ala
130 135 140

Val Gly Ala Leu Glu Ala Gln Leu Lys Leu Lys Thr Gly Lys Leu Val
145 150 155 160

Ser Leu Ser Ala Gln Asn Leu Val Asp Cys Ser Thr Glu Lys Tyr Gly
165 170 175

Asn Lys Gly Cys Asn Gly Gly Phe Met Thr Thr Ala Phe Gln Tyr Ile
180 185 190

Ile Asp Asn Lys Gly Ile Asp Ser Asp Ala Ser Tyr Pro Tyr Lys Ala
195 200 205

Met Asp Leu Lys Cys Gln Tyr Asp Ser Lys Tyr Arg Ala Ala Thr Cys
210 215 220

Ser Lys Tyr Thr Glu Leu Pro Tyr Gly Arg Glu Asp Val Leu Lys Glu
225 230 235 240

Ala Val Ala Asn Lys Gly Pro Val Ser Val Gly Val Asp Ala Arg His
245 250 255

Pro Ser Phe Phe Leu Tyr Arg Ser Gly Val Tyr Tyr Glu Pro Ser Cys
260 265 270

Thr Gln Asn Val Asn His Gly Val Leu Val Val Gly Tyr Gly Asp Leu
275 280 285

Asn Gly Lys Glu Tyr Trp Leu Val Lys Asn Ser Trp Gly His Asn Phe

290

295

300

Gly Glu Glu Gly Tyr Ile Arg Met Ala Arg Asn Lys Gly Asn His Cys
305 310 315 320

Gly Ile Ala Ser Phe Pro Ser Tyr Pro Glu Ile
325 330

<210> 2
<211> 576
<212> DNA
<213> Homo sapiens

<220>
<221> Human cDNA sequence for Cathepsin S
<222> (1)...(576)
<223>

<400> 2
gctttcagtg ctgtggggc cctggaagca cagctgaagc tgaaaacagg aaagctggtg 60
tctctcagtg cccagaacct ggtggattgc tcaactgaaa aatatggaaa caaaggctgc 120
aatggtggtc tcatgacaac ggcttccag tacatcattg ataacaaggg catcgactca 180
gacgcttcct atccctacaa agccatggat ctgaaatgtc aatatgactc aaaatatcgt 240
gctgccacat gttcaaagta cactgaactt ccttatggca gagaagatgt cctgaaagaa 300
gctgtggcca ataaaggccc agtgtctgtt ggtgttagatg cgcgcatcc ttctttcttc 360
ctctacagaa gtggtgtcta ctatgaacca tcctgtactc agaatgtgaa tcatggtgta 420
cttgggtttg gctatggtga tcttaatggg aaagaatact ggcttgtgaa aaacagctgg 480
ggccacaact ttggtaaga aggatatatt cggatggcaa gaaataaagg aaatcattgt 540
gggattgcta gcttccttc ttacccagaa atctag 576

<210> 3
<211> 225
<212> PRT
<213> Homo sapiens

<220>
<221> Amino acid sequence for residues 114-331 of Cathepsin S
<222> (1)...(225)
<223> with the addition of a C-terminal Glycine and six histidine tag

<400> 3

Ile Leu Pro Asp Ser Val Asp Trp Arg Glu Lys Gly Cys Val Thr Glu

1

5

10

15

Val Lys Tyr Gln Gly Ser Cys Gly Ala Cys Trp Ala Phe Ser Ala Val
20 25 30

Gly Ala Leu Glu Ala Gln Leu Lys Leu Lys Thr Gly Lys Leu Val Ser
35 40 45

Leu Ser Ala Gln Asn Leu Val Asp Cys Ser Thr Glu Lys Tyr Gly Asn
50 55 60

Lys Gly Cys Asn Gly Gly Phe Met Thr Thr Ala Phe Gln Tyr Ile Ile
65 70 75 80

Asp Asn Lys Gly Ile Asp Ser Asp Ala Ser Tyr Pro Tyr Lys Ala Met
85 90 95

Asp Gln Lys Cys Gln Tyr Asp Ser Lys Tyr Arg Ala Ala Thr Cys Ser
100 105 110

Lys Tyr Thr Glu Leu Pro Tyr Gly Arg Glu Asp Val Leu Lys Glu Ala
115 120 125

Val Ala Asn Lys Gly Pro Val Ser Val Gly Val Asp Ala Arg His Pro
130 135 140

Ser Phe Phe Leu Tyr Arg Ser Gly Val Tyr Tyr Glu Pro Ser Cys Thr
145 150 155 160

Gln Asn Val Asn His Gly Val Leu Val Val Gly Tyr Gly Asp Leu Asn
165 170 175

Gly Lys Glu Tyr Trp Leu Val Lys Asn Ser Trp Gly His Asn Phe Gly
180 185 190

Glu Glu Gly Tyr Ile Arg Met Ala Arg Asn Lys Gly Asn His Cys Gly
195 200 205

Ile Ala Ser Phe Pro Ser Tyr Pro Glu Ile Gly His His His His His
210 215 220

His
225

<210> 4
<211> 340
<212> PRT
<213> Homo sapiens

<220>
<221> Amino acid sequence for residues 1-331 of Cathepsin S
<222> (3)..(333)
<223>

<220>
<221> Amino acid sequence for residues 1-331 of Cathepsin S
<222> (3)..(333)
<223> Additional N-terminal Methionine-Proline and a C-terminal Glycine
- 6x-histidine tag

<400> 4

Met Pro Met Lys Arg Leu Val Cys Val Leu Leu Val Cys Ser Ser Ala
1 5. 10 15

Val Ala Gln Leu His Lys Asp Pro Thr Leu Asp His His Trp His Leu
20 25 30

Trp Lys Lys Thr Tyr Gly Lys Gln Tyr Lys Glu Lys Asn Glu Glu Ala
35 40 45

Val Arg Arg Leu Ile Trp Glu Lys Asn Leu Lys Phe Val Met Leu His
50 55 60

Asn Leu Glu His Ser Met Gly Met His Ser Tyr Asp Leu Gly Met Asn
65 70 75 80

His Leu Gly Asp Met Thr Ser Glu Glu Val Met Ser Leu Met Ser Ser
85 90 95

Leu Arg Val Pro Ser Gln Trp Gln Arg Asn Ile Thr Tyr Lys Ser Asn
100 105 110

Pro Asn Arg Ile Leu Pro Asp Ser Val Asp Trp Arg Glu Lys Gly Cys
115 120 125

Val Thr Glu Val Lys Tyr Gln Gly Ser Cys Gly Ala Cys Trp Ala Phe
130 135 140

Ser Ala Val Gly Ala Leu Glu Ala Gln Leu Lys Leu Lys Thr Gly Lys
145 150 155 160

Leu Val Ser Leu Ser Ala Gln Asn Leu Val Asp Cys Ser Thr Glu Lys
165 170 175

Tyr Gly Asn Lys Gly Cys Asn Gly Gly Phe Met Thr Thr Ala Phe Gln
180 185 190

Tyr Ile Ile Asp Asn Lys Gly Ile Asp Ser Asp Ala Ser Tyr Pro Tyr
195 200 205

Lys Ala Met Asp Gln Lys Cys Gln Tyr Asp Ser Lys Tyr Arg Ala Ala
210 215 220

Thr Cys Ser Lys Tyr Thr Glu Leu Pro Tyr Gly Arg Glu Asp Val Leu
225 230 235 240

Lys Glu Ala Val Ala Asn Lys Gly Pro Val Ser Val Gly Val Asp Ala
245 250 255

Arg His Pro Ser Phe Phe Leu Tyr Arg Ser Gly Val Tyr Tyr Glu Pro
260 265 270

Ser Cys Thr Gln Asn Val Asn His Gly Val Leu Val Val Gly Tyr Gly
275 280 285

Asp Leu Asn Gly Lys Glu Tyr Trp Leu Val Lys Asn Ser Trp Gly His
290 295 300

Asn Phe Gly Glu Glu Gly Tyr Ile Arg Met Ala Arg Asn Lys Gly Asn
305 310 315 320

His Cys Gly Ile Ala Ser Phe Pro Ser Tyr Pro Glu Ile Gly His His
325 330 335

His His His His
340